

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: June 25, 2003, 11:43:10 ; Search time 36.6 Seconds

(without alignments)  
84,446 Million cell updates/sec

Title: US-09-869-540A-2\_COPY\_5\_19  
Perfect score: 90  
Sequence: 1 LRCMLGRVYRPMQV 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

SPTREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mmc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriophage:\*  
17: sp\_archaeophages:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	100.0	77	6 Q9TTS8	Q9TTS8 sus scrofa
2	90	100.0	165	4 Q8WVG0	Q8WVG0 mus musculus
3	90	100.0	165	11 Q9D220	Q9D220 mus musculus
4	81	90.0	150	13 Q91916	Q91916 paratuberculin
5	81	85.6	86	4 Q9BQD1	Q9BQD1 homo sapien
6	72	80.0	86	4 Q9BQD1	Q9BQD1 homo sapien
7	45.5	50.6	85	11 Q88259	Q88259 mus musculus
8	44.5	49.4	166	16 Q8X2X6	Q8X2X6 escherichia
9	44.5	48.9	154	11 Q9D3R4	Q9D3R4 mus musculus
10	43.5	48.3	397	5 Q76862	Q76862 drosophila
11	43.5	48.3	414	5 Q9W4M4	Q9W4M4 drosophila
12	42	46.7	103	6 Q951L0	Q951L0 macaca fasc
13	42	46.7	250	16 P71972	P71972 mycobacteri
14	42	46.7	311	5 Q9GS12	Q9GS12 caenorhabdi
15	42	46.7	355	12 Q91HR4	Q91HR4 pseudorabd
16	42	46.7	358	12 Q91HR5	Q91HR5 pseudorabd

17	42	46.7	577	12 Q9Q3F7	Q9Q3F7 pseudorabd
18	42	46.7	578	12 Q9PYC0	Q9PYC0 pseudorabd
19	42	46.7	1489	16 Q8B9P5	Q8B9P5 agrobacteri
20	42	46.7	2749	10 Q8SA93	Q8SA93 zea mays (m
21	41	45.6	304	11 Q9R165	Q9R165 mus musculu
22	41	45.6	331	12 Q91TU4	Q91TU4 lupula harp
23	41	45.6	345	17 Q8TUC2	Q8TUC2 methanosarc
24	41	45.6	362	5 Q9VCX6	Q9VCX6 drosophila
25	41	45.6	610	12 Q8VBR8	Q8VBR8 pea seed-bo
26	41	45.6	621	16 Q8XSF4	Q8XSF4 raietonia s
27	41	45.6	633	5 Q9V754	Q9V754 drosophila
28	41	45.6	3198	12 Q91W34	Q91W34 pea seed-bo
29	41	45.6	3199	12 Q85074	Q85074 pea seed-bo
30	40.5	45.0	173	16 Q8A1B5	Q8A1B5 streptococc
31	40.5	45.0	524	11 Q8H1L0	Q8H1L0 mus musculu
32	40.5	45.0	640	11 Q8VBM1	Q8VBM1 mus musculu
33	40.5	45.0	732	4 Q13032	Q13032 homo sapien
34	40	44.4	124	16 Q8XQZ9	Q8XQZ9 raietonia s
35	40	44.4	346	12 Q9DMG8	Q9DMG8 rat cytoleg
36	40	44.4	448	5 Q9V4C0	Q9V4C0 drosophila
37	40	44.4	448	5 Q9V4C0	Q9V4C0 drosophila
38	40	44.4	480	16 P71091	P71091 bacillus su
39	40	44.4	493	4 Q8TBS2	Q8TBS2 homo sapien
40	40	44.4	493	11 P70539	P70539 raietonia s
41	40	44.4	493	11 P70603	P70603 raietonia s
42	40	44.4	493	11 P70603	P70603 raietonia s
43	40	44.4	1131	5 Q02222	Q02222 caenorhabdi
44	40	44.4	1705	11 Q9ERK5	Q9ERK5 mus musculu
45	39.5	43.9	65	6 Q29326	Q29326 sus scrofa

## ALIGNMENTS

RESULT 1	ID	Q9TTS8	PRELIMINARY:	PRT:	77 AA.
AC	Q9TTS8	01-MAY-2000 (TREMREL. 13, Created)			
DT	01-MAY-2000 (TREMREL. 13, Last sequence update)				
DT	01-JUN-2002 (TREMREL. 21, Last annotation update)				
DE	Melanin-concentrating hormone (fragment).				
OS	Sus scrofa (pig).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.				
OX	NCBI_TaxID:9623;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	TISSUE-HYPOTHALAMUS;				
RT	Mattari R.L.; Dyer C.J.;				
RT	"Partial cDNA sequence of porcine melanin-concentrating hormone				
RT	(MCH)."				
RL	Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.				
DR	EMBL: AF203034; AF09587.1; -				
DR	PRINTS; PRO1641; PROMCHFAMILY.				
FT	NON_TER	1			
FT	CHAIN	59	>77		
FT	NON_TER	77			
SO	SEQUENCE	77 AA;	8861 MW;	CB640E22E4CF3F21 CRC64;	
Query Match					
Best Local Similarity		100.0%;	Score 90;	DB 6;	Length 77;
Matches	15;	Conservative	0;	Mismatches	0;
Indels	0;	Gaps	0;		
DB	1 LRCMLGRVYRPMQV 15				
QY					
DB	63 LRCMLGRVYRPMQV 77				
RESULT 2					
Q8WVG0					
ID	Q8WVG0	PRELIMINARY:	PRT:	165 AA.	
AC	Q8WVG0;				

```

DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical 18.7 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC018048; AAI18048.1; -
DR PRINTS: PR01641; PROMCHFAMILY.
KW Hypothetical protein.
SQ SEQUENCE 165 AA; 18679 MW; 2339E8938BEC499 CRC64;

Query Match 100.0%; Score 90; DB 4; Length 165;
Best Local Similarity 100.0%; Pred. No. 7,8e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRCMLGRVYRRCMOV 15
DB 151 LRCMLGRVYRRCMOV 165

RESULT 3
09D20
ID 09D20 PRELIMINARY; PRT: 165 AA.
AC 09D20;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DE A230109K23R1K protein.
GN A230109K23R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=HYPOPHALAMUS;
RA MEDLINE-21085660; Pubmed-11217851;
RA Kawai J., Shimagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yamada I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schiml L.M., Staudt F., Suzuki K., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.F.,
RA Guernicich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzaretti P.,
RA Norone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Saeki H., Sato K., Schoenbach C., Sessa T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyshna-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT *Functional annotation of a full-length mouse cDNA collection.*;
RL Nature 409:685-690(2001).
DR EMBL: AK020723; BAB32189.1; -
DR MGD: MGI:1925014; A230109K23R1K.
DR PRINTS: PR01641; PROMCHFAMILY.
SQ SEQUENCE 165 AA; 18516 MW; 1872B6B1D4BEEAC2 CRC64;

Query Match 100.0%; Score 90; DB 11; Length 165;
Best Local Similarity 100.0%; Pred. No. 7,8e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 LRCMLGRVYRRCMOV 15
DB 151 LRCMLGRVYRRCMOV 165

RESULT 4
09I9L6
ID 09I9L6 PRELIMINARY; PRT: 150 AA.
AC 09I9L6;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Melanin-concentrating hormone-like protein.
OS Paratichthys olivaceus (Flounder).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphi; Acanthopterygii; Percomorphi; Pleuronectiformes;
OC Pleuronectoidae; Paratichthyidae; Paratichthys.
OX NCBI_TaxID=8255;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Jeon J., Lee J., Song Y.;
RT *Melanin-concentrating hormone-like protein.*;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF236090; AAF67166.1; -
DR PRINTS: PR01641; PROMCHFAMILY.
SQ SEQUENCE 150 AA; 16913 MW; 117A280F46ED4499 CRC64;

Query Match 90.0%; Score 81; DB 13; Length 150;
Best Local Similarity 85.7%; Pred. No. 2,5e-06;
Matches 12; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRCMLGRVYRRCMOV 14
DB 132 LRCMLGRVYRRCMOV 145

RESULT 5
09B0D1
ID 09B0D1 PRELIMINARY; PRT: 86 AA.
AC 09B0D1;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DE Pro-melanin-concentrating hormone-like 2 protein.
GN PCHL2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-TESTIS;
RA MEDLINE-21108351; Pubmed-11181993;
RA Courseaux A., Nahon J.L.;
RT *Birth of two chimeric genes in the Homnidae lineage.*;
RL Science 281:1293-1297(2001).
DR EMBL: AY008413; AAK31297.1; -
DR EMBL: AY028320; AAK31290.1; -
DR EMBL: AY008412; AAK31296.1; -
DR PRINTS: PR01641; PROMCHFAMILY.
SQ SEQUENCE 86 AA; 9856 MW; 1C8D698786B47471 CRC64;

Query Match 85.6%; Score 77; DB 4; Length 86;
Best Local Similarity 86.7%; Pred. No. 7,2e-06;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

## RESULT 6

ID	ORGID	PRELIMINARY	PRT	86 AA.
DT	09B010			
DT	09B010			
DT	01-JUN-2001 (TREMBLrel. 17, Created)			
DT	01-JUN-2001 (TREMBLrel. 17, Last sequence update)			
DE	01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE	Pro-melanin-concentrating hormone-like 1 protein.			
GN	PMCHL1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=TESTIS;			
RX	MEDLINE=21108351; PubMed=1181993;			
RA	Courseaux A., Nahon J.L.			
RT	"Birth of two chimeric genes in the Homnidae lineage."			
RL	Science 291:1293-1297(2001).			
DR	EMBL; AY008411; AAK31295.1; -			
DR	EMBL; AY028318; AAK31289.1; -			
DR	PRINTS; PR01641; PROMCHFAMILY.			
SQ	SEQUENCE 86 AA; 9715 MW; BC5EC470EDBA867F3 CRC64;			
Query Match	80.0%;	Score 72;	DB 4;	Length 86;
Best Local Similarity	80.0%;	Pred. No. 5.2e-05;		
Matches 12;	Conservative 1;	Mismatches 2;	Indels 0;	Gaps 0;
OY	1 LRCLGRVYRPMQV 15			
db	72 LSCMLGRVYQSCWQV 86			

## RESULT 7

ID	088259	PRELIMINARY:	PRF:	85 AA.
AC	088259;			
DT	01-NOV-1998 (TREMblrel. 08, Created)			
DT	01-NOV-1998 (TREMblrel. 08, last sequence update)			
DT	01-MAR-2002 (TREMblrel. 20, last annotation update)			
DE	Maz554 (Fragment).			
GN	MSZF54.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
NCBI_OX	NCBI_TaxID=10090;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57Bl6.			
RX	MEDLINE=98296253; PubMed=9630514;			
RA	Agata I., Matsuda E., Shimizu A.,			
RT	Rapid and efficient cloning of cDNAs encoding Krueppel-1like zinc			
RL	finger proteins by degenerate PCR."			
RL	Gene 213:55-64(1998).			
DR	EMBL; AB010364; BAA31420.1; -			
DR	HSSP; P08048; ZNF.			
DR	InterPro: IPR000822; Znf.CZH2.			
DR	Pfam: PF00096; zfc2h2.3			
DR	SMART: SM00355; ZNF_CZH2_3			
DR	PROSITE; PS00028; ZINC_FINGER_CZH2_1; 3.			
DR	PROSITE; PS0157; ZINC_FINGER_CZH2_2; 3.			
KW	DNA-binding; metal-binding; zinc-finger.			
FT	NON_TER	1		
FT	NON_TER	85		
FT	NON_TER	85		
SEQUENCE	85 AA; 10007 MW; F0FA94CA59DBDEDA CRC64;			
Query Match	50.6%; Score 45.5; DB 11; Length 85;			
Best Local Similarity	46.7%; Pred. NO.1,8;			
Matches	7; Conservative 5; Mismatches 2; Indels 1; Gaps 1;			
2 RC-MLRVTRPCMOV	5			
1	1			

Db 33 KCEVCGKVERTCWL 47

## RESULT 8

ID	08X2X6	PRELIMINARY;	PRT;	186 AA.
AC	08X2X6;			
DT	01-MAR-2002 (TREMBLrel. 20, Created)			
DT	01-MAR-2002 (TREMBLrel. 20, last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, last annotation update)			
DE	Hypothetical protein EC64988.			
GN	EC64988.			
OS	Escherichia coli O157:H7.			
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;			
OC	Escherichia.			
OX	NCBI_TaxID=83334;			
OX	(1)			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-O157:H7 / RIMD 0509952;			
RX	MEDLINE=211556233; Pubmed=11258796;			
RX	Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,			
RA	Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,			
RA	Iida T., Takemi H., Honda T., Sasakiwa C., Ogasawara N., Yasunaga T.,			
RA	Kuhara S., Shiba T., Hattori M., Shimagawa H.;			
RT	*Complete genome sequence of enterohemorrhagic Escherichia coli			
RT	O157:H7 and genomic comparison with a laboratory strain K-12.*			
RL	DNA Res. 8:11-23(2001)			
DR	EMBL: AP002567; BAB38411.1; -			
KM	Hypothetical protein.			
SO	SEQUENCE 186 AA; 21454 MW; CSEDI1B6BCA9898 CRC64;			

**Best L**

Matches	Conservative	Mismatches	Indels	Gaps
Qy	1			
	LRMCGRVYRPMCQ 14			
Db	159			
	LECILNR-YKPAHQ 171			
RESULT 9				
ID	Q9DJR4	PRELIMINARY;	PRT;	154 AA.
AC	Q9DJR4;			
DT	01-JUN-2001 (TREMBLrel, 17, Created)			
DT	01-JUN-2001 (TREMBLrel, 17, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel, 17, Last annotation update)			
DE	4933438J11RLK protein.			
GN	4933438J11RLK.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
NCBI_OX	NCBI_TaxID=10090;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=TESTIS;			
RC	MEDLINE=210855660; PubMed=11217851;			
RA	Kawai J, Shingawa A, Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,			
RA	Alizawa K., Iizawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,			
RA	Salto T., Ohtsuki Y., Gojibori T., Bono H., Kasukawa T., Salto R.,			
RA	Kadota K., Matzuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,			
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,			
RA	Schirral L.M., Stahlb J., Suzuki R., Tomita M., Wagner L., Mashio T.,			
RA	Sakai K., Okido T., Furuno M., Isono H., Baldairelli R., Barch G.,			
RA	Blake J., Boffelli D., Bojunga N., Carrincci P., de Bonaldi M.F.,			
RA	Brownstein M.J., Bult C., Flechner C., Fujita M., Gariboldi M.F.,			
RA	Guslinich S., Hill D., Hofmann M., Hume D.A., Kamlya M., Lee N.H.,			
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mommaerts P.,			
RA	Nordone P., Ring B., Ringwald C., Sessa T., Shibata Y., Storch K.-F.,			
RA	Sasaki H., Sato K., Schoenbach C., Sessa T., Shibata Y., Storch K.-F.,			
RA	Suzuki H., Toyochika K., Wang K.H., Wetz C., Whitaker C., Wilmink L.,			

RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RT Nature 409:685-690(2001).  
 DR EMBL: AK017122; BAB30609.1;  
 DR MGI:1918557; 4933439J11Rik.  
 SO SEQUENCE 154 AA; 17921 MW; 91B9DCAC58434C49 CRC64;

Query Match 48.9%; Score 44; DB 11; Length 154;  
 Best Local Similarity 61.5%; Pred. No. 5.9;  
 Matches 8; Conservative 1; Mismatches 2; Indels 2; Gaps 1;  
 Oy 3 CML-GRRYRCW 13  
 Db 127 CLCTKREKPCW 139

RESULT 10  
 ID 076862 PRELIMINARY; PRT; 397 AA.  
 AC 076862;  
 DT 01-NOV-1998 (TREMBlrel. 08, Created)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
 DE EG:100G10.2 protein.  
 DE EG:100G10.2 OR CG2601.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 RN NCBI\_TaxID:7227;  
 RP SEQUENCE FROM N.A.  
 RA Ferraz C., Vidal S., Brun C., Bucheton A., Demaille J.G.;  
 RT "Sequencing the distal X chromosome of Drosophila melanogaster";  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RA Benos P.;  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL023874; CAI19647.1;  
 DR Flybase: FBgn0024997; EG:100G10.2.  
 DR InterPro: IPR004162; Slna.  
 DR Pfam: PF03145; Slna; 1  
 SO SEQUENCE 397 AA; 44818 MW; 98AA331AF8494565 CRC64;

Query Match 48.3%; Score 43.5; DB 5; Length 397;  
 Best Local Similarity 46.7%; Pred. No. 18;  
 Matches 7; Conservative 4; Mismatches 3; Indels 1; Gaps 1;  
 Oy 1 LRCLMGRRYRCW-WQ 14  
 Db 62 MKCFMGRRYRCWQ 76

RESULT 11  
 ID 0944M4 PRELIMINARY; PRT; 414 AA.  
 AC 0944M4;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DE EG:100G10.2 protein.  
 DE EG:100G10.2 OR CG2681.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 RN NCBI\_TaxID:7227;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RX MEDLINE:20196006; PubMed-10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blair R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miles G.L.G.,  
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Plankhach C., Baldwin D.,  
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,  
 RA Borovda D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jaitani M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclio J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reimert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M., Smith T.,  
 RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svrtkars R., Teeter C., Turner R., Venter E., Wang A., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RT Science 287:2185-2195(2000).  
 DR EMBL: AE003425; AAF45811.1;  
 DR Flybase: FBgn0024997; EG:100G10.2.  
 DR InterPro: IPR004162; Slna.  
 DR Pfam: PF03145; Slna; 1  
 SO SEQUENCE 414 AA; 46719 MW; 0F70A32A08B7640 CRC64;

Query Match 48.3%; Score 43.5; DB 5; Length 414;  
 Best Local Similarity 46.7%; Pred. No. 19;  
 Matches 7; Conservative 4; Mismatches 3; Indels 1; Gaps 1;  
 Oy 1 LRCLMGRRYRCW-WQ 14  
 Db 79 MKCFMGRRYRCWQ 93

RESULT 12  
 ID 0951L0 PRELIMINARY; PRT; 103 AA.  
 AC 0951L0;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DE EG:100G10.2 protein.  
 DE EG:100G10.2 OR CG2681.  
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecoidea; Macaca.  
 RN NCBI\_TaxID:9541;  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-TESTIS;  
 RA Hashimoto K., Osada N., Hida M., Kusuda J., Tanuma R., Hirai M.,  
 RA Terao K., Sugano S.;  
 RT "Isolation of novel full-length cDNA clones from macaque testis cDNA  
 libraries";

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB072778; BAB69747.1;  
 KW Hypothetical protein.  
 SQ SEQUENCE 103 AA; 11326 MW; D8547BDDA141195 CRC64;

Query Match 46.7%; Score 42; DB 6; Length 103;  
 Best Local Similarity 46.2%; Pred. No. 8;  
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 3 CMLGRVYRPMQV 15  
 DB 52 CDLGILSNPCWRL 64

## RESULT 13

P71972 PRELIMINARY; PRT; 250 AA.

AC P71972;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Hypothetical protein RV2675c.  
 GN RV2675C OR MTCY441.44C OR MT2749.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteriales;  
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-H37RV;  
 RX MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekela F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborn J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Sulten G., Taylor K., Whitehead S., Barrell B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence."  
 RL Nature 393:537-544(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CDC 1551 / OSHKOSH;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Helt D., Hickey E.,  
 RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
 RA Bishai W.;  
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains."  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: STRONG. IN C-TERMINAL REGION TO M.LEPRAE U1764Z.  
 DR EMBL: Z80225; CAB0328.1;  
 DR EMBL: AE007105; AKK47064.1;  
 DR TIGR: MT2749;  
 DR Tuberculin: RV2675c;  
 DR InterPro: IPR000051; SAM\_bind.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 250 AA; 27545 MW; 2EC9718C7550F32C CRC64;

Query Match 46.7%; Score 42; DB 16; Length 250;  
 Best Local Similarity 33.3%; Pred. No. 21;  
 Matches 9; Conservative 4; Mismatches 2; Indels 12; Gaps 1;

OY 1 LRCML-----GRVYRPMQV 15  
 DB 218 VRCVLERFAIKPMLVGRVHAPMEV 244

## RESULT 14

O9GS12

ID O9GS12 PRELIMINARY; PRT; 311 AA.

AC O9GS12;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE F56H6.13 protein.  
 GN F56H6.13.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kershaw J.K.;  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology."  
 RL Science 282:2012-2018(1998).  
 DR EMBL: Z81553; CAC14353.1;  
 SQ SEQUENCE 311 AA; 36614 MW; 4A9960CE41C1027B CRC64;

Query Match 46.7%; Score 42; DB 5; Length 311;  
 Best Local Similarity 33.3%; Pred. No. 25;  
 Matches 5; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 1 LRCMLGRVYRPMQV 15  
 DB 155 MRCVRLYINGIMNI 169

## RESULT 15

O91HR4 PRELIMINARY; PRT; 555 AA.

AC O91HR4;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Glycoprotein E (Fragment).  
 OS Pseudorabies virus.  
 OC Alphaherpesvirinae; Varicelloviridae;  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OX NCBI\_TaxID=10345;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GNANDONG;  
 RA Lou G.M.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF403050; AKK95640.1;  
 DR InterPro: IPR003404; Herpes-glycopE.  
 DR Pfam: PF02480; Herpes\_gE; 1.  
 FT NON\_TER 1  
 FT NON\_TER 555  
 FT NON\_TER 555  
 SQ SEQUENCE 555 AA; 59871 MW; 79CBEF2F03003CDB CRC64;

Query Match 46.7%; Score 42; DB 12; Length 555;  
 Best Local Similarity 63.6%; Pred. No. 44;  
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 2 RCMLGRVYRPMQV 12  
 DB 252 RCLLYVYEP 262

Search completed: June 25, 2003, 11:54:03  
 Job time : 37.6 secs